FDA's Draft Guidance for Industry

Non -Inferiority Clinical Trials

Some key statistical issues and concepts

Robert T. O'Neill Ph.D.

Director , Office of Biostatistics

OTS,CDER, FDA

For presentation at the 4th Annual FDA/DIA Statistics Forum: Integrating Knowledge In Clinical Development: Meta-Analysis, Non-Inferiority, and Related Topics; April 19-21, 2010

NI Guidance - Contents

- Part 1
  - A general discussion of regulatory, study design, scientific, and statistical issues associated with the use of non-inferiority studies when these are used to establish the effectiveness of a new drug.
- Part 2
  - Focuses on the quantitative analytical and statistical approaches used to determine the non-inferiority margin for use in NI studies, as well as the advantages and disadvantages of available analysis methods.

NI Guidance Contents

- Part 3
  - Addresses some questions about NI studies and provides practical advice to relevant regulatory issues (Q & A)
- Part 4
  - Includes four examples of successful and unsuccessful efforts to define non-inferiority margins and to conduct NI studies. - that is, when a NI study will not provide the evidence needed

Part 2 - Key Features

- Statistical approach to estimation and selection of the margin M and to the statistical analysis of the NI study results
- NI study provides two comparisons
  - Direct comparison of test and active comparator
  - Indirect comparison of the test with (unobserved) placebo

Choosing a metric for treatment effect and margin selection

- Absolute difference in cure rates
- Relative measures: relative risk, risk ratio,hazard ratio, odds ratio
- Log of relative or ratio measures

- Understanding how the magnitude of the margin relates to the clinical benefit

Metrics of treatment effect and comparison

- Log (Hazard ratio) or Log(relative risk) is preferred
- Symmetry of margin
  - Log (A/B) = - Log(B/A)
- Statistical properties
- Ease of understanding

Sources of uncertainty

- In the quantification of the treatment effect of the active control
- In the NI study , itself and its analysis

Three main sources of uncertainty in the conclusions from a NI study

1) Precision of the active control effect size and its heterogeneity , if any
   - Point estimates not acceptable, confidence interval needed
   - Estimate of study to study variability, lack of consistent effect size
   - Choice of confidence limits (95%, 90%, 85%)
2) Constancy assumption: the need to assume that the effect size from the historical active control studies will not change in the NI study
3) Risk of making a wrong decision from the test of hypothesis in the NI study - The type 1 error - always controlled at 0.025 one sided

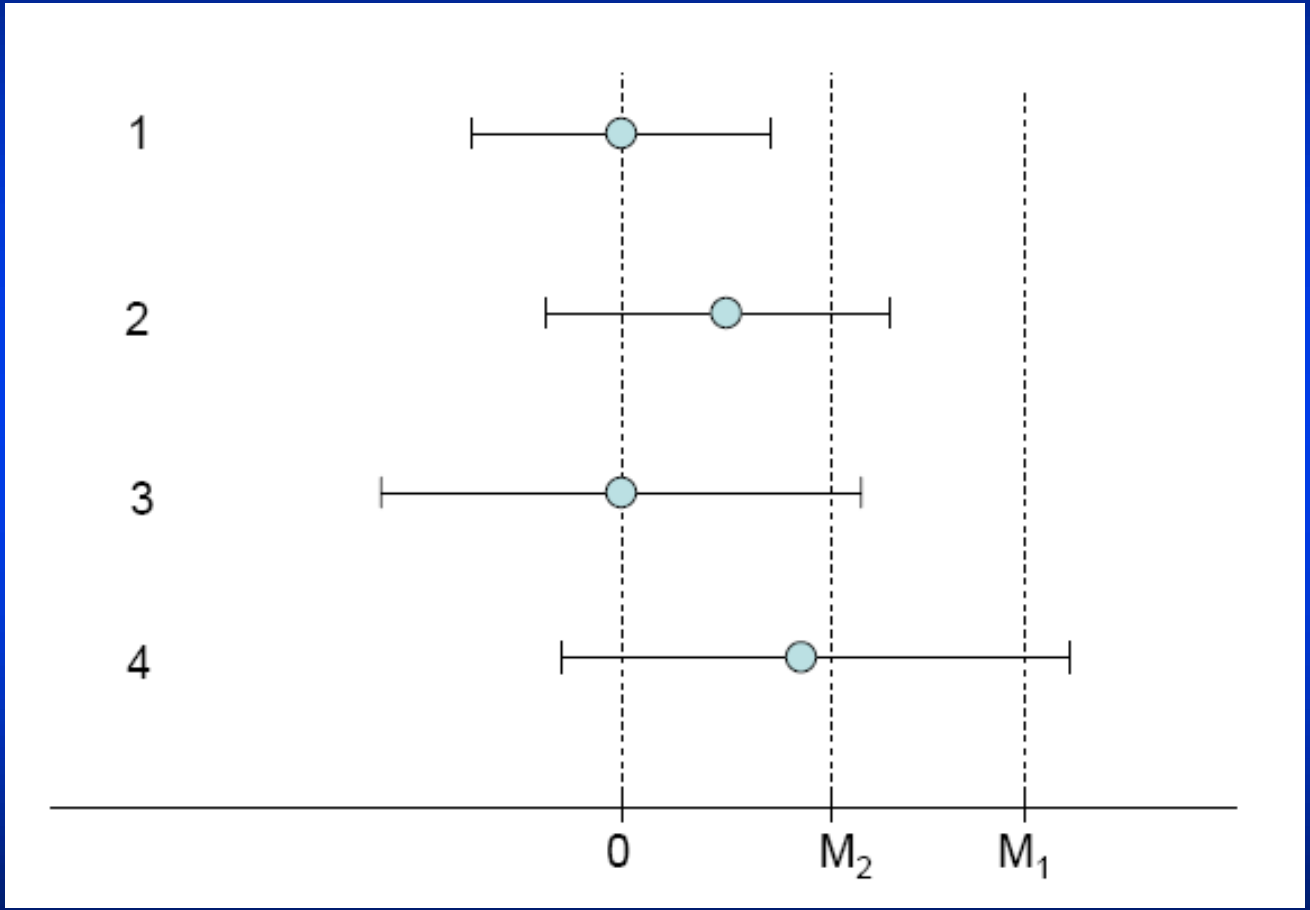
Quantification of treatment effect of the active comparator

- From a single study : 99% confidence interval ?
- From multiple studies: use a random effects models for effect size and confidence interval
  - Excessive heterogeneity, especially consistent with no effect of active control is an issue
- When possible, identify the underpowered studies from studies truly with no effect
- Use all studies - drop none, except with good reason

Discussion of the fixed margin and the synthesis method

- Fixed margin approach also called the two confidence interval approach
- Synthesis approach combines data and its variability (from historical placebo controlled studies) to infer a (non randomized) comparison with the unobserved placebo under the constancy assumption

Figure 3. Active Control – Test Drug differences
Point estimate, 95% CI



Control – Test (C-T)
(degree of inferiority of test drug)

Interpretation

1. C-T point estimate = 0 and upper bound of 95% CI < M_2, indicating test drug is effective (NI demonstrated).
2. Point estimate of C-T favors C and upper bound of 95% CI < M_1 but > M_2, indicating effect > 0 but unacceptable loss of the control effect.
3. Point estimate of C-T is zero and upper bound of 95% CI < M_1 but it is slightly greater than M_2. Judgment could lead to conclusion of effectiveness.
4. C-T point estimate favors C and upper bound of 95% CI > M_1, indicating there is no evidence of effectiveness for test drug.

Fixed margin and Synthesis Approach

- A discussion of what each is
- A discussion of pros and cons
- Illustration of the approaches on the same data set
- A recommendation

Which method to use ?

- Advice is to use fixed margin for ruling out M_1 . Why ? This is the efficacy claim - forces the specific estimate of the active control effect size - conservative approach is intended
- May use the synthesis method for M_2 demonstration . Why ? Assumptions and judgment more relevant and flexible
  - May be some unusual pathological examples

Explaining the synthesis method to laymen

- Where the apparent efficiency comes from
- Explaining that the type 1 error control is not the same so the methods are not directly comparable
- Asking the question of whether the test drug would have been superior to placebo when it is not part of the randomized comparison in the NI study introduces potential confusion
- Percent preservation of the active comparator effect (without explicit M_1)

The Concept of 'Discounting' M_1

- To adjust estimate of M for uncertainties in data and conditions of NI study
- Not always done, and not always very conservative - should be reasonable

Concept of Discounting the margin size for lack of constancy and other issues

- Generally , only M_1 is discounted and before M_2 settled
- M_2 , since it is a function (percent) of M_1 , is not discounted
- Caveat: after the NI study is completed and if the conduct of the study was such to make the constancy assumption suspect, further discounting may be appropriate: how much to discount is challenging - might lend itself to regression adjustments !
- Concept of adjusting for the distribution of the effect modifiers that are different in the NI study when contrasted with the historical data because the mix of the patient populations differs in a meaningful way

Type 1 error control for the NI study

- Synthesis and Fixed Margin are not controlling the same error

Type 1 errors are not directly comparable or interpreted the same

While this graph may suggest that the synthesis test is always statistically more efficient (i.e., requiring smaller sample size) than the fixed margin approach, it is difficult to compare the two methods in any meaningful sense because each method uses different statistical error rates , or Type 1 errors. This fact needs to be considered in the interpretation of the conclusions of each method. The fixed margin method is controlling a type 1 error rate within the NI study that is conditional on the pre-specified fixed NI margin chosen. The synthesis method is controlling an unconditional error rate for the null hypothesis it is testing (imagining that all the studies could be repeated infinitely often), but it does so assuming the constancy assumption is satisfied.

Taken from example 1(a)

the numerator. In this case, the estimated log-Relative Risk of ximelegatran versus warfarin is 0.329 (log of 1.39) with a standard error of 0.216 while the estimated log-Relative Risk of warfarin versus placebo is -1.02 (log of .527) with a standard error of 0.154. The estimate of the log warfarin effect is -1.02, and the standard error of this estimate is 0.154; these estimates are combined with the NI data as if all the data were in a randomized comparison with placebo. The synthesis test statistic is calculated as:

Fixed Margin approach:

Conclude non-inferiority if the upper limit of the 95% CI for the log-hazard ratio of the experimental treatment to the active control is less than ½ of the lower limit of the 95% CI for the log-hazard ratio of placebo to the active control.

$$\hat{\gamma}_{T/C} + 1.96\sqrt{Var\left\{\hat{\gamma}_{T/C}\right\}} < \frac{1}{2}\left\{\hat{\gamma}_{P/C} - 1.96\sqrt{Var\left\{\hat{\gamma}_{P/C}\right\}}\right\}$$

$$\hat{\gamma}_{T/C} + 1.96\sqrt{Var\left\{\hat{\gamma}_{T/C}\right\}} < -\frac{1}{2}\hat{\gamma}_{C/P} - 1.96\sqrt{\frac{1}{4}Var\left\{\hat{\gamma}_{C/P}\right\}}$$

$$\frac{\hat{\gamma}_{T/C} + \frac{1}{2}\hat{\gamma}_{C/P}}{\sqrt{Var\left\{\hat{\gamma}_{T/C}\right\}} + \sqrt{\frac{1}{4}Var\left\{\hat{\gamma}_{C/P}\right\}}} < -1.96$$

Synthesis test

$$\frac{\hat{\gamma}_{T/C} + \frac{1}{2}\hat{\gamma}_{C/P}}{\sqrt{Var\left\{\hat{\gamma}_{T/C}\right\} + \frac{1}{4}Var\left\{\hat{\gamma}_{C/P}\right\}}} < -1.96$$

The Claimed Efficiency of the Fixed vs. Synthesis Methods It comes from the standard error expression

$$\frac{0.329 + \frac{1}{2}\{-1.02\}}{\sqrt{0.216^2 + \left\{\frac{1}{2}\{0.154\}\right\}^2}} < -1.96$$

Synthesis method

$$\frac{0.329 + \frac{1}{2}\{-1.02\}}{0.216 + \frac{1}{2}\{0.154\}} < -1.96$$

Fixed Margin

Appendix - Examples

- Five examples illustrate different aspect of choosing a NI margin, or the application of a method of NI analysis, and other considerations relevant to whether it is possible to conduct and interpret the results of a NI study

The objective of the Examples

- All very different issues and approaches
  - One size does not fit all
- Illustrate the application of the logic of the guidance to real examples
- Illustrate application of methods and interpretation

EX. 1
Determination of an NI margin of a new anticoagulant - Fixed Margin approach

- How to select and assess the randomized trials of the active control on which to base the estimate of active comparator treatment effect
- How to assess whether the assumption of assay sensitivity is appropriate, and whether the constancy assumption is reasonable for this drug class
- Why it is appropriate to use a conservative (e.g. 95% lower bound) for estimating the treatment effect size of the active comparator, accounting for between study variability, and considering other uncertainties in the randomized trial data
- The use of the lower bound of the 95% confidence interval in the NI study for C-T to demonstrate non-inferiority

Ex 1(b): Application of the Synthesis Method to Ex 1(a)

- Critical features of the synthesis approach to demonstrating the NI of a new anticoagulant
- The calculations and sources of statistical variability that are incorporated in the synthesis approach
- The main differences in interpretation of the fixed margin and the synthesis approaches when applied to the same set of studies and data
  - Fixed margin approach for M_1, the synthesis approach for M_2

Ex. 2: Determination of a Non-inferiority margin for complicated urinary tract infection (cUTI) - Fixed Margin Approach

- Use of the absolute difference in cure rates as the metric of treatment effect
- Determination of a NI margin when there are no randomized active comparator placebo-controlled studies available for the indication of interest (cUTI)
- Estimating the placebo response rate in cUTI based upon data from uncomplicated urinary tract infections (generally a less severe form of urinary tract infection leading to a high (conservative) estimate
- Importance of seeking out all relevant studies for margin determination and incorporating the limitations of the studies, the analyses and the resulting estimates in the consideration of the estimate of the NI margin
- The approach is credible only when the effect size is large

Ex. 3: Aspirin to prevent death or death/MI after myocardial infarction

- When it may not be possible to determine the NI margin because of the limitations of the data available

Table 7. Results of six placebo-controlled randomized studies of the effect of aspirin in secondary prevention of death or MI after myocardial infarction

Study	Year published	Aspirin		Placebo		Relative Risk (95% CI)
		N	Event rate	N	Event rate	
MRC-1	1974	615	9.9%	624	13.1%	0.75 (0.55, 1.03)
CDP	1976	758	9.5%	771	12.5%	0.76 (0.57, 1.02)
MRC-2	1979	832	16.0%	850	22.2%	0.72 (0.59, 0.88)
GASP	1978	317	13.6%	309	17.5%	0.78 (0.54, 1.12)
PARIS	1980	810	17.4%	406	22.7%	0.77 (0.61, 0.97)
AMIS	1980	2267	18.6%	2257	19.2%	0.97 (0.86, 1.09)

*the event rate of either group needs further verification from each article

- The NI margin based on these six studies is so small that a trial to rule out loss of this effect would be unrealistically large. For mortality, the largest and most recent trial showed no significant effect.

Ex 4: Xeloda to treat metastatic colorectal cancer - the synthesis method

- The use of the synthesis method to demonstrate a loss of no more than 50% of the historical control treatment's effect and a relaxation of this criterion when two NI are available
- The use of supportive endpoints in the decision making process
- The use of a conservative estimate of the control treatment effect size, because a subset of the available studies to estimate the margin was selected and the effect was measured relative to a previous standard of care instead of placebo

This is the time to send in comments on the draft

The docket is open for comments

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.